#### SEQUENCE LISTING

#### (1) GENERAL INFORMATION

- (i) APPLICANT:
   Seidel, Christoph; Weinhues, Ursula-Henrike;
   Schmitt, Urban; Motz, Manfred; Wiedmann, Michael;
   Upmeier, Barbara; Soutscheck, Erwin
- (ii) TITLE OF INVENTION:

  Recombinant antigen from the NS3 region of the hepatitis C virus
- (iii) NUMBER OF SEQUENCES: 9
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Felfe & Lynch
  - (B) STREET: 805 Third Avenue
  - (C) CITY: New York
  - (D) STATE: New York
  - (E) COUNTRY: USA
  - (F) ZIP: 10022
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
  - (B) COMPUTER: IBM PS/2
  - (C) OPERATING SYSTEM: PC-DOS
  - (D) SOFTWARE: PatentIn Relase #1.0, Version #1.25 (EPA)
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: 08/511,759
  - (B) FILING DATE: 7-AUGUST-1995y
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: p 44 28 705.4
  - (B) FILING DATE: 12 AUGUST 1994
- (viii) ATTORNEY/AGENT INFORMATION
  - (A) NAME: HANSON, NORMAN
  - (B) REGISTRATION NUMBER: 30, 946
  - (C) REFERENCE/DOCKET NUMBER: HUBR 1067.1
- (ix) TELECOMMUNICATION INFORMATION
  - (A) TELEPHONE: 212-688-9200
  - (B) TELEFAX: 212-838-3884

#### INFORMATION FOR SEQ ID NO: (2) SEQUENCE CHARACTERISTICS: (i) LENGTH: 885 base pairs (A) TYPE: nucleic acid (B) STRANDEDNESS: both (C) TOPOLOGY: linear (D) (ii) MOLECULE TYPE: cDNA (vi) INITIAL ORIGIN: (A) ORGANISM: hepatitis C virus POSITION IN GENOME: (viii) CHROMOSOME SEGMENT: (A) (ix) CHARACTERISTICS: (A) NAME/KEY: CDS LOCATION: 1..885 (B) (xi) SEQUENCE DESCRIPTION: SEO ID NO: 1: ATG ACC ATG ATT ACG AAT TCC CGG GGA TCC ATC ATG AAA TCC CCG GTG 48 Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val TTC ACG GAT AAC TCC TCT CCA CCG GTA GTG CCC CAG AGC TTC CAG GTG 96 Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val 20 GCT CAC CTG CAT GCT CCC ACA GGC AGC GGC AAG AGC ACC AAG GTC CCG 144 Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro GCT GCA TAC GCA GCT CAG GGC TAC AAG GTG CTA GTG CTC AAC CCT TCT 192 Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser 50 GTT GCT GCA ACA TTG GGC TTT GGT GCC TAC ATG TCC AAG GCT CAT GGG 240 Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly 80 65 ATC GAT CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC 288 Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser 90 CCC ATT ACG TAC TCC ACT TAC GGC AAG TTT CTT GCC GAC GGC GGG TGC 336 Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys 105 100

-	,				ţ	*								٠.			
			GCT Ala														384
			TCC Ser														432
			GCG Ala														480
			GTG Val														528
			ATC Ile 180													·	576
⊈lys Ū			AGA Arg														624
∭ ∭GAG ∏Glu □																	672
U TAC Tyr 225	CGC Arg	GGT Gly	CTT Leu	GAC Asp	GTG Val 230	TCC Ser	GTC Val	ATC Ile	CCG Pro	ACC Thr 235	AGC Ser	GGT Gly	GAT Asp	GTT Val	GTC Val 240		720
GTC Val	GTG Val	GCA Ala	ACC Thr	GAC Asp 245	GCC Ala	CTC Leu	ATG Met	ACC Thr	GGC Gly 250	TAT Tyr	ACC Thr	GGC Gly	GAC Asp	TTC Phe 255	GAC Asp	•	768
HTCG	GTG Val	ATA Ile	GAC Asp 260	TGC Cys	AAC Asn	ACG Thr	TGT Cys	GTC Val 265	ACT Thr	CAG Gln	ACA Thr	GTC Val	GAT Asp 270	TTC Phe	AGC Ser		816
			ACC Thr														864
			ACT Thr														885
				41													

### (3) <u>INFORMATION FOR SEQ ID NO:</u> 2

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 295 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val 1 5 10 15

Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val 20 25 30

Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro 35 40 45

ন্ৰAla Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser ত্ৰ 50 60

Wal Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly
65 70 75 80

Tile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser

Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys

Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr 120 125

Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu 130 135 140

Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly
145 150 155 160-

Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr 165 170 175

Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile 180 185 190

Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp 195 200 205

Leu 210	Ala	Thr	Lys	Leu	Val 215	Ala	Met	Gly	Ile	Asn 220	Ala	Val	Ala	Tyr
Arg	Gly	Leu	Asp	Val 230	Ser	Val	Ile	Pro	Thr 235	Ser	Gly	Asp	Val	Val 240
Val	Ala	Thr	Asp 245	Ala	Leu	Met	Thr	Gly 250	Tyr	Thr	Gly	Asp	Phe 255	Asp
Val	Ile	Asp 260	Cys	Asn	Thr	Cys	Val 265	Thr	Gln	Thr	Val	Asp 270	Phe	Ser
Asp	Pro 275	Thr	Phe	Thr	Ile	Glu 280	Thr	Thr	Thr	Leu	Pro 285	Gln	Asp	Ala
Ser 290	Arg	Thr	Gln	Arg	Arg 295									
	210 Arg Val Val Asp Ser	Arg Gly  Val Ala  Val Ile  Asp Pro 275  Ser Arg	Arg Gly Leu  Val Ala Thr  Val Ile Asp 260  Asp Pro Thr 275  Ser Arg Thr	Arg Gly Leu Asp  Val Ala Thr Asp 245  Val Ile Asp Cys 260  Asp Pro Thr Phe 275  Ser Arg Thr Gln	Arg Gly Leu Asp Val 230  Val Ala Thr Asp Ala 245  Val Ile Asp Cys Asn 260  Asp Pro Thr Phe Thr 275  Ser Arg Thr Gln Arg	Arg Gly Leu Asp Val Ser 230  Val Ala Thr Asp Ala Leu 245  Val Ile Asp Cys Asn Thr 260  Asp Pro Thr Phe Thr Ile 275  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val 230  Val Ala Thr Asp Ala Leu Met 245  Val Ile Asp Cys Asn Thr Cys 260  Asp Pro Thr Phe Thr Ile Glu 280  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile 230  Val Ala Thr Asp Ala Leu Met Thr 245  Val Ile Asp Cys Asn Thr Cys Val 265  Asp Pro Thr Phe Thr Ile Glu Thr 275  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile Pro 230  Val Ala Thr Asp Ala Leu Met Thr Gly 250  Val Ile Asp Cys Asn Thr Cys Val Thr 265  Asp Pro Thr Phe Thr Ile Glu Thr Thr 275  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile Pro Thr 235  Val Ala Thr Asp Ala Leu Met Thr Gly Tyr 250  Val Ile Asp Cys Asn Thr Cys Val Thr Gln 265  Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr 275  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser 235  Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr 250  Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr 260  Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu 275  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly 230 Val Ala Thr Asp Ala Leu Met Thr Gly 250 Tyr Thr Gly Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val 260 Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro 275 Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Val Ile Asp 260 Cys Asn Thr Cys Val Thr Gln Thr Val Asp 270  Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln 285  Ser Arg Thr Gln Arg Arg	Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe 255  Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe 270  Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ser Arg Thr Gln Arg Arg

- (4) <u>INFORMATION FOR SEQ ID NO</u>: 3:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 40 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRANDEDNESS: single
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGGGATCCA TCATGAAATC CCCGGTGTTC ACGGATAACT

(5) <u>INFORMATION FOR SEQ ID NO:</u> 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC

40

(6)	INFO	RMATION FOR SEO ID NO: 5:	
	(i)	SEQUENCE CHARACTERISTICS:  (A) LENGTH: 39 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: cDNA	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 5:	
GAGG	GATCC.	A TCATGAAAGC GGTGGACTTT ATCCCTGTG	39
(7)	INFO	RMATION FOR SEQ ID NO: 6:	
	(i)	SEQUENCE CHARACTERISTICS:  (A) LENGTH: 33 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: cDNA	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 6:	
GAGA	AGCTT'	T TAACACGTGT TGCAGTCTAT CAC	33
[ (8) D	INFO	RMATION FOR SEQ ID NO: 7:	
	(i)	SEQUENCE CHARACTERISTICS:  (A) LENGTH: 39 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	
	(ii)	MOLECULE TYPE: cDNA	
	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
GAGG	GATCC	A TCATGAAACA CCTGCATGCT CCCACCGGC	39

#### (8) INFORMATION FOR SEQ ID NO:

- (i) SEQUENCE CHARACTERISTICS:
  - LENGTH: 33 base pairs (A)
  - TYPE: nucleic acid (B)
  - (C) STRANDEDNESS: single
  - TOPOLOGY: linear (D)
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

#### GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC

- (3) INFORMATION FOR SEQ ID NO:
  - (i) SEQUENCE CHARACTERISTICS:
    - LENGTH: 302 amino acids (A)
    - (B) TYPE: amino acid
    - (D) TOPOLOGY: linear
  - (ii) MOLECULE TYPE: protein

 $\Box$ 

Ū <u>o</u>

(xi) SEQUENCE DESCRIPTION:

□Pro Val Glu Asn Leu Glu Thr Thr Met Arg Ser Pro Val Phe Thr □ 1 5 10 15

🕮 Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val Ala

His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro

Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro

Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala

His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr

Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala 95 100

The sources

					•									
Asp	Gly	Gly	Cys	Ala 110	Gly	Gly	Ala	Tyr	Asp 115	Ile	Ile	Ile	Cys	Asp 120
Glu	Cys	His	Ser	Thr 125	Asp	Ala	Thr	Ser	Ile 130	Leu	Gly	Ile	Gly	Thr 135
Val	Leu	Asp	Gln	Gly 140	Glu	Thr	Ala	Gly	Ala 145	Lys	Leu	Val	Val	Phe 150
Ala	Thr	Ala	Thr	Pro 155	Pro	Gly	Ser	Val	Thr 160	Val	Pro	His	Pro	Asn 165
Ile	Glu	Glu	Val	Ala 170	Leu	Ser	Thr	Thr	Gly 175	Glu	Ile	Pro	Phe	Tyr 180
Gly	Lys	Ala	Ile	Pro 185	Leu	Glu	Val	Ile	Lys 190	Gly	Gly	Arg	His	Leu 195
Ile	Phe	Cys	His	Ser 200	Lys	Arg	Lys	Cys	Asp 205	Glu	Leu	Ala	Thr	Lys 210
	Val	Ala	Met	Gly 215	Ile	Asn	Ala	Val	Ala 220	Tyr	Tyr	Arg	Gly	Leu 225
型 <b>基</b> sp 型	Val	Ser	Val	Ile 230	Pro	Thr	Ser	Gly	Asp 235	Val	Val	Val	Val	Ala 240
o Thr	_			245					250					255
∏lle □	Asp	Cys	Asn	Thr 260	Cys	Val	Thr	Gln	Thr 265	Val	Asp	Phe	Ser	Leu 270
Map Map	Pro	Thr	Phe	Thr 275	Ile	Glu	Thr	Thr	Thr 280	Leu	Pro	Gln	Asp	Ala 285
	Ser	Arg	Thr	Gln 290	Arg	Arg	Gly	Arg	Thr 295	Gly	Arg	Gly	Lys	Pro 300
Gly	Ile													

Gly Ile 302

#### SEQUENCE PROTOCOL

- (1) GENERAL INFORMATION:
  - (i) APPLICANT:
    - (A) NAME: Boehringer Mannheim GmbH
    - (B) ROAD: Sandhofer Str. 112-132
    - (C) CITY: Mannheim
    - (E) COUNTRY: Germany
    - (F) POSTAL CODE: 68305
  - (ii) TITLE OF APPLICATION: Recombinant antigen from the NS3 region of the hepatitis C virus
  - (iii) NUMBER OF SEQUENCES: 8
    - (iv) COMPUTER READABLE FORM:
      - (A) DATA CARRIER: Floppy disk
      - (B) COMPUTER: IBM PC compatible
      - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
      - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPA)
- (2) INFORMATION FOR SEQ ID NO: 1:
  - (i) SEQUENCE CHARACTERISTICS:
    - (A) LENGTH: 885 base pairs
    - (B) TYPE: nucleic acid
    - (C) STRAND FORM: both
    - (D) TOPOLOGY: linear
  - (ii) TYPE OF MOLECULE: cDNA

### (vi) INITIAL ORIGIN:

(A) ORGANISM: hepatitis C virus

# (viii) POSITION IN THE GENOME:

(A) CHROMOSOME/SEGMENT: NS3

## (ix) CHARACTERISTICS:

(A) NAME/KEY: CDS

(B) LOCATION: 1..885

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG Met U 1	ACC Thr	ATG Met	ATT	ACG Thr 5	AAT Asn	TCC Ser	CGG Arg	GGA Gly	TCC Ser 10	Ile	ATG Met	AAA Lys	TCC	CCG Pro 15	GTG Val	48
配 即 即	ACG Thr	GAT Asp	AAC Asn 20	Ser	TCT Ser	CCA Pro	CCG Pro	GTA Val 25	GTG Val	CCC Pro	CAG Gln	AGC Ser	TTC Phe 30	CAG Gln	GTG Val	96
GCT Ala	CAC His	CTG Leu 35	CAT His	GCT Ala	CCC Pro	ACA Thr	GGC Gly 40	AGC Ser	GGC Gly	AAG Lys	AGC Ser	ACC Thr 45	AAG Lys	GTC Val	CCG Pro	144
GCT Ala	GCA Ala 50	TAC Tyr	GCA Ala	GCT Ala	CAG Gln	GGC Gly 55	TAC Tyr	AAG Lys	GTG Val	CTA Leu	GTG Val 60	CTC Leu	AAC Asn	CCT Pro	TCT Ser	192
GTT	GCT Ala	GCA Ala	ACA Thr	TTG Leu	GGC Gly 70	TTT Phe	GGT Gly	GCC Ala	TAC Tyr	ATG Met 75	TCC Ser	AAG Lys	GCT Ala	CAT His	GGG Gly 80	240
ATC	GAT	CCT	AAC	ATC	AGG	ACC	GGG	GTG	AGA	ACA	ATT	ACC	ACT	GGC	AGC	288
Ile	Asp	Pro	Asn	Ile 85	Arg	Thr	Gly	Val	Arg 90	Thr	Ile	Thr	Thr	Gly 95	Ser	
CCC Pro	ATT Ile	ACG Thr	TAC Tyr 100	TCC Ser	ACT Thr	TAC Tyr	GGC Gly	AAG Lys 105	TTT	CTT Leu	GCC Ala	GAC Asp	GGC Gly 110	GGG Gly	TGC Cys	336
GCA Ala	GGG Gly	GGT Gly 115	GCT Ala	TAT Tyr	GAC Asp	ATA Ile	ATA Ile 120	ATT Ile	TGT Cys	GAC Asp	GAG Glu	TGC Cys 125	CAC His	TCC Ser	ACG Thr	384
GAT Asp	GCC Ala 130	ACA Thr	TCC Ser	ATC <sup>'</sup> Ile	TTG Leu	GGC Gly 135	ATC Ile	GGC Gly	ACT Thr	GTC Val	CTT Leu 140	GAC Asp	CAA Gln	GGA Gly	GAG Glu	432

ÁCT Thr 145	GCG Ala	GGG Gly	GCG Ala	AAA' Lys	Leu 150	GTT Val	GTG Val	TTC Phe	GCC Ala	ACC Thr 155	GCĆ Ala	ec Thr	CCT Pro	CCG Pro	GGC Gly 160	480
TCC Ser	GTC Val	ACT Thr	GTG Val	CCC Pro 165	CAT His	CCC Pro	AAC Asn	ATT Ile	GAG Glu 170	GAG Glu	GTT Val	GCT Ala	CTA Leu	TCC Ser 175	ACC Thr	528
						TAC Tyr										576
						ATC Ile										624
						GTC Val 215										672
						TCC Ser										720
GTC Val	GTG Val	GCA Ala	ACC Thr	GAC Asp 245	GCC Ala	CTC Leu	ATG Met	ACC Thr	GGC Gly 250	TAT Tyr	ACC Thr	GGC Gly	GAC Asp	TTC Phe 255	GAC Asp	768
TCG Ser	GTG Val	ATA Ile	GAC Asp 260	TGC Cys	AAC Asn	ACG Thr	TGT Cys	GTC Val 265	ACT Thr	CAG Gln	ACA Thr	GTC Val	GAT Asp 270	TTC Phe	AGC Ser	816
CTT Ieu	GAC Asp	CCT Pro 275	ACC Thr	TTC Phe	ACC Thr	ATT	GAG Glu 280	ACG Thr	ACC Thr	ACA Thr	CTT Leu	CCC Pro 285	CAG Gln	GAT Asp	GCT Ala	864
GTC	TCC Ser 290	CGC Arg	ACT Thr	CAA Gln	CGA	CGG Arg 295										885

### (2) INFORMATION FOR SEQ ID NO: 2:

### (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

### (ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

/207

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val

1 5 10 15

Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val

20 25 30

Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro 35 40 45

Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser 50 60

Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly 65 75 80

Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser 85 90 95

Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys 100 105 110

Ala Gly Gly Ala Tyr Asp Ile Ile Cys Asp Glu Cys His Ser Thr 115 120 125

Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu 130 135 140

Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly 145 150 155 160

Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr 165 170 175

Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile 180 185 190

Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp 195 200 205

Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr 210 215 220

Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val 225 230 235 240

Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp 245 250 255

Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser 260 265 270

Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala 275 280 285

Val Ser Arg Thr Gln Arg Arg 290 295

	•
(2) INFORMATION FOR SEQ ID NO: 3:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 40 base pairs	
(B) TYPE: nucleic acid	
(C) STRAND FORM: single	
(D) TOPOLOGY: linear	
(ii) TYPE OF MOLECULE: CDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
AAGGGATCCA TCATGAAATC CCCGGTGTTC ACGGATAACT	40
(2) INFORMATION FOR SEQ ID NO: 4:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 39 base pairs	
(B) TYPE: nucleic acid	
(C) STRAND FORM: single	
(D) TOPOLOGY: linear	
(ii) TYPE OF MOLECULE: cDNA	·
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4	
GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC	39
(2) INFORMATION FOR SEQ ID NO: 5:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 39 base pairs	
(B) TYPE: nucleic acid	
(C) STRAND FORM: singl	

(ii) TYPE OF MOLECULE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5	
GAGGGATCCA TCATGAAAGC GGTGGACTTT ATCCCTGTG	<b>39</b>
(2) INFORMATION FOR SEQ ID NO: 6:	
(i) SEQUENCE CHARACTERISTICS:	Ψ
(A) LENGTH: 33 base pairs	
(B) TYPE: nucleic acid	
(C) STRAND FORM: single	
(D) TOPOLOGY: linear	
(ii) TYPE OF MOLECULE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6	
GAGAAGCTTT TAACACGTGT TGCAGTCTAT CAC	33
(2) INFORMATION FOR SEQ ID NO: 7:	•
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 39 base pairs	·
(B) TYPE: nucleic acid	
(C) STRAND FORM: single	
(D) TOPOLOGY: linear	

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
GAGGGATCCA TCATGAAACA CCTGCATGCT CCCACCGGC	39
(2) INFORMATION FOR SEQ ID NO: 8:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 33 base pairs	
(B) TYPE: nucleic acid	
(C) STRAND FORM: single	
(D) TOPOLOGY: linear	
(ii) TYPE OF MOLECULE: cDNA	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8	
GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC	33